**Name: Saloni Surendra Alshi**

**Assignment 2: Basic Statistics\_Level-2**

**SET 2**

**Topics: Normal distribution, Functions of Random Variables**

1. The time required for servicing transmissions is normally distributed with *μ* = 45 minutes and *σ* = 8 minutes. The service manager plans to have work begin on the transmission of a customer’s car 10 minutes after the car is dropped off and the customer is told that the car will be ready within 1 hour from drop-off. What is the probability that the service manager cannot meet his commitment?
2. 0.3875
3. 0.2676
4. 0.5
5. 0.6987

**Answer:**

**Using R-studio:**

> # let 'X' be the time required for servicing transmissions.

> # 'X' is normally distributed with μ = 45 minutes and σ = 8 minutes.

> # i.e X~N(μ=45, σ(sq) =64)

> # The service manager plans to have work begin on the transmission of a customer’s

> # car 10 minutes after the car is dropped off and the customer is told that the

> # car will be ready within 1 hour from drop-off

> # so he has only 50 minutes left to complete his work

> # To find the probability that the service manager cannot meet his commitment,

> #i.e we have to find P(X>50)

> # therefore, P[X>z-score]

> # =P[X>(x-μ)/σ]

> prob=1-pnorm(50,45,8)

> prob

[1] 0.2659855

Therefore, the probability that the service manager cannot meet his commitment is 0.2659855 which is approximately equal to 0.2676 so the option B is correct answer.

**(Done in python)**

1. The current age (in years) of 400 clerical employees at an insurance claims processing center is normally distributed with mean *μ* = 38 and Standard deviation *σ* =6. For each statement below, please specify True/False. If false, briefly explain why.
2. More employees at the processing center are older than 44 than between 38 and 44.

**Answer:**

Using R-studio,

> # X~N (μ=38, σ(sq)=36 )

> #A) Finding prob.that the processing center are older than 44 than between 38 and 44

> # i.e P[38<x<44], P[x>44]

>probability=1-pnorm(44,38,6)

>probability

[1] 0.1586553

i.e 63 employees out of 400

> prob=pnorm(44,38,6)-pnorm(38,38,6)

> prob

[1] 0.3413447

i.e 34.13447% \*400=136.53788=137 employees out of 400

therefore, more employees at the processing center are older than 44 than between 38 and 44.

This statement is false.

1. A training program for employees under the age of 30 at the center would be expected to attract about 36 employees.

**Answer:**

**Using R-studio,**

> #B) A training program for employees under the age of 30 at the center would

> #be expected to attract about 36 employees i.e P[x<=30]

> probability=pnorm(30,38,6)

> probability

[1] 0.09121122

i.e 9.12%= 36 employees out of 400

therefore, A training program for employees under the age of 30 at the center would be expected to attract about 36 employees.

This statement is true.

**Also, done in python.**

1. If *X1* ~ *N*(μ, σ2) and *X*2 ~ *N*(μ, σ2) are *iid* normal random variables, then what is the difference between 2 *X*1 and *X*1 + *X*2? Discuss both their distributions and parameters.

**Answer:**

If *X1* ~ *N*(μ, σ2) and *X*2 ~ *N*(μ, σ2) are *iid* normal random variables, then

1. *2X1* ~ *N*(2μ, 4σ2) According to the product property of normal distribution
2. *X*1 + *X*2 **~** N (μ+ μ, σ(sq)+ σ(sq)) According to the addition property of normal dist.

*i.e X*1 + *X*2 **~** N (2μ, 2σ(sq))

Difference between these two is,

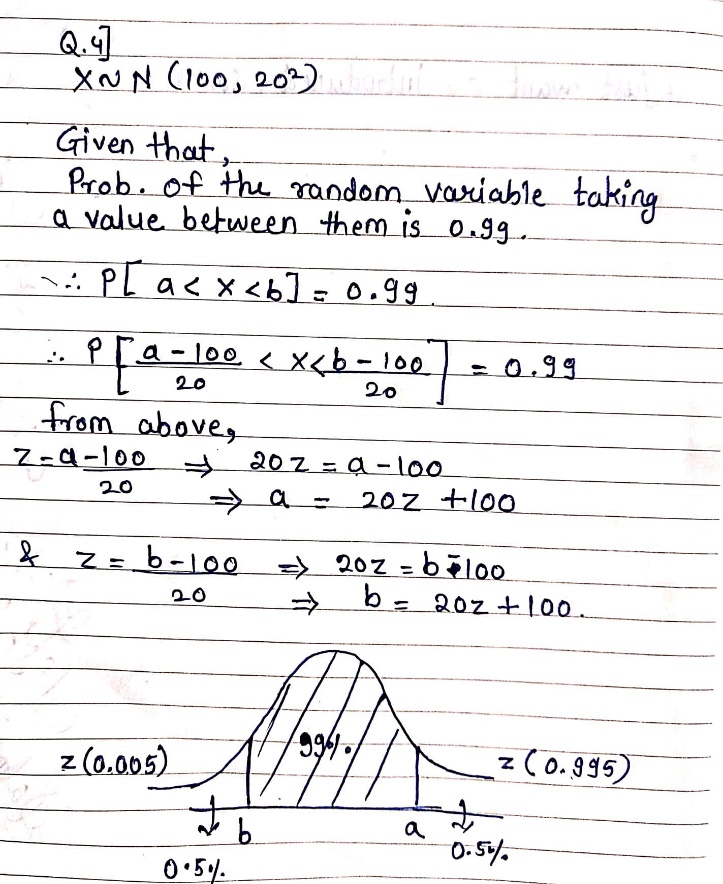
*2X1* – (*X*1 + *X*2) ~ N (0, 6 σ(sq)) According to the difference property of normal dist.

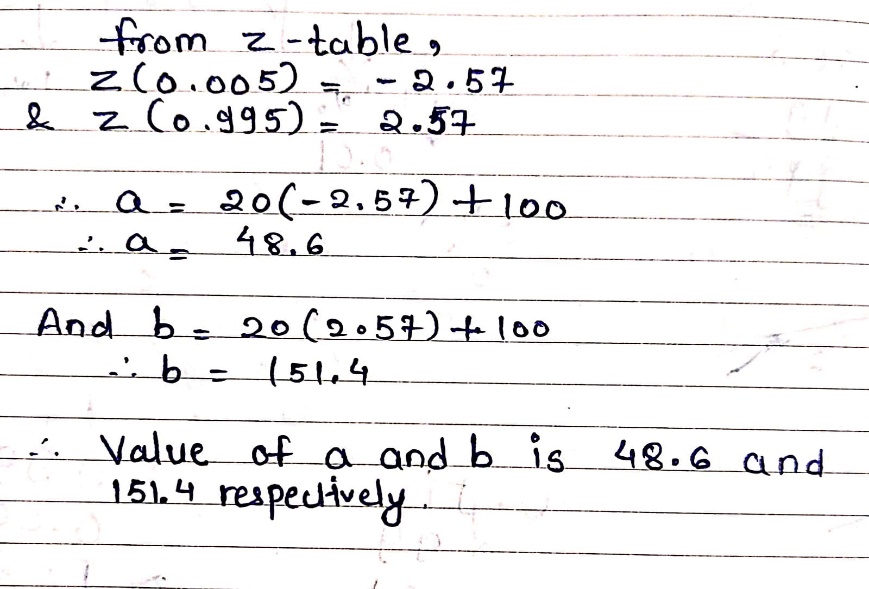
The mean of 2X1and X1+X2 is same but the var of 2X1 is 2 times more than the variance of X1+X2.The difference between the two says that the two given variables are **identically** and **independently** distributed.

1. Let X ~ N(100, 202). Find two values, *a* and *b*, symmetric about the mean, such that the probability of the random variable taking a value between them is 0.99.
2. 90.5, 105.9
3. 80.2, 119.8
4. 22, 78
5. 48.5, 151.5
6. 90.1, 109.9

**Answer:**

**Manually,**

****



**Using R-studio,**

> # finding z score value

> round(qnorm(0.005),3)

[1] -2.576

> round(qnorm(0.995),3)

[1] 2.576

> a=20\*(-2.575829)+100

> a

[1] 48.48342

> b=20\*(2.575829)+100

> b

[1] 151.5166

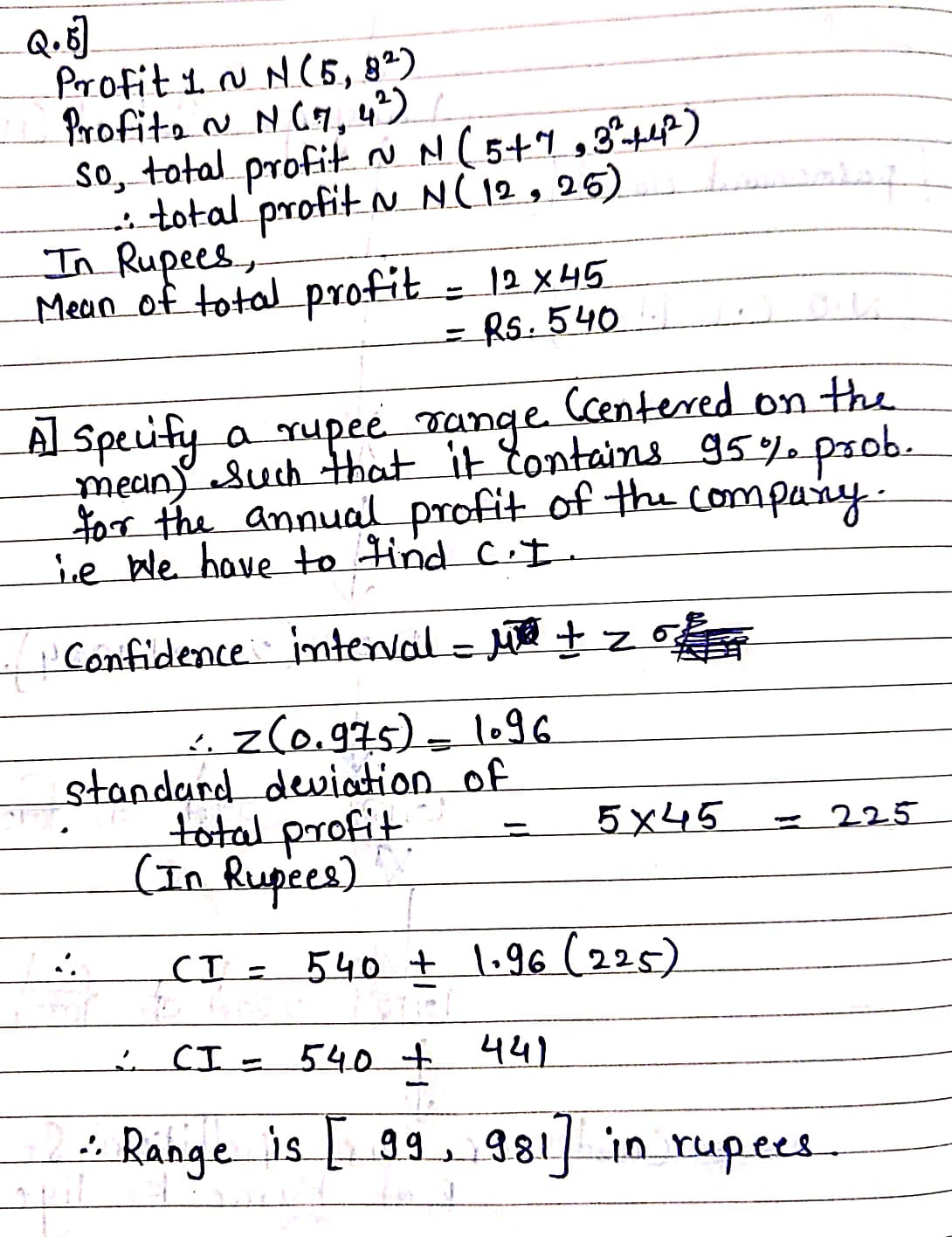
Correct option is D.

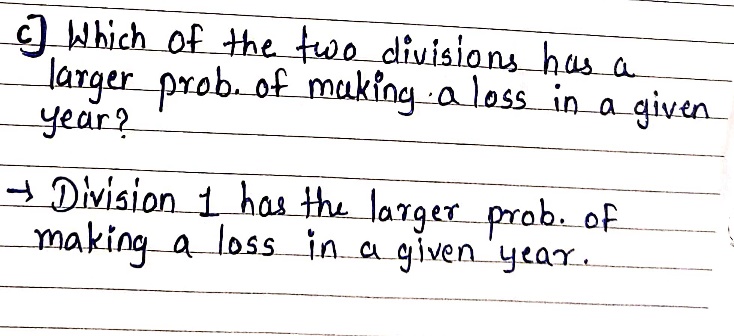
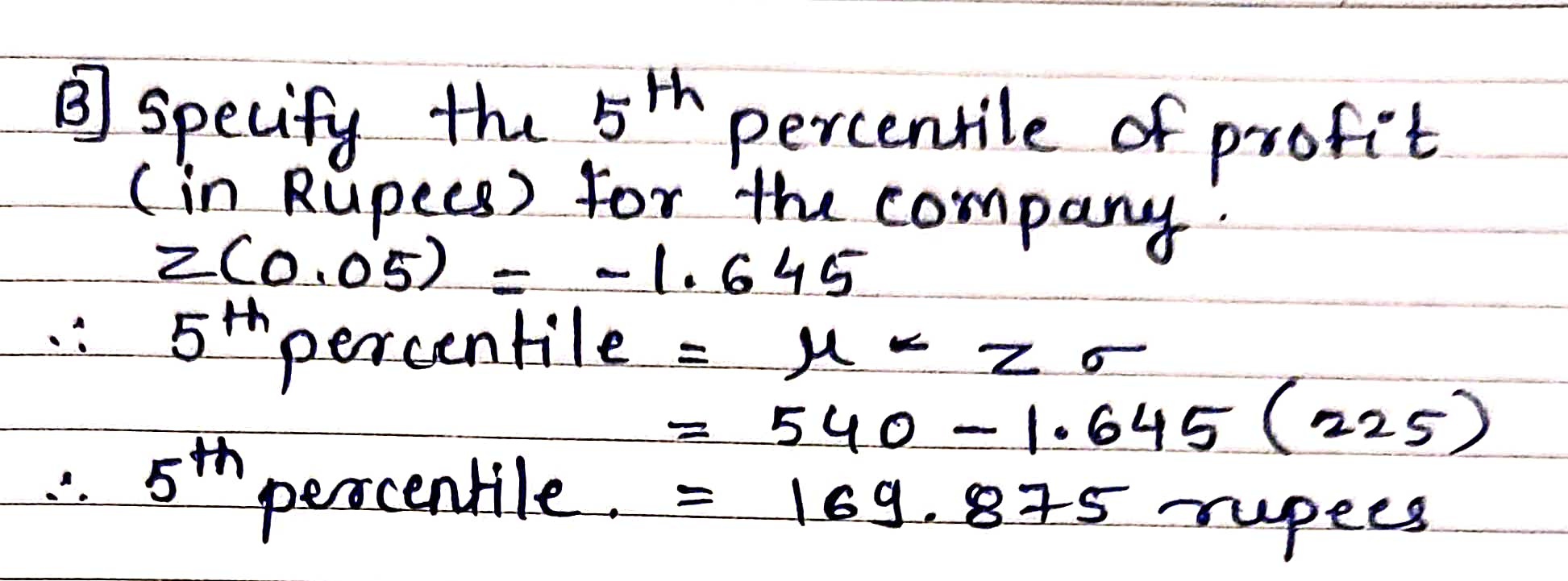
**Also, done in python.**

1. Consider a company that has two different divisions. The annual profits from the two divisions are independent and have distributions Profit1 ~ N(5, 32) and Profit2 ~ N(7, 42) respectively. Both the profits are in $ Million. Answer the following questions about the total profit of the company in Rupees. Assume that $1 = Rs. 45
2. Specify a Rupee range (centered on the mean) such that it contains 95% probability for the annual profit of the company.
3. Specify the 5th percentile of profit (in Rupees) for the company
4. Which of the two divisions has a larger probability of making a loss in a given year?

**Answer:**

**Manually,**

****

****

**Using R-studio,**

> # profit1~N(5,9)

> # profit2~N(7,16)

> # total\_profit~N(12,25) #according to addition property of normal distribution

> # In rupees,

> mean\_of\_totalprofit=12\*45

> mean\_of\_totalprofit

[1] 540

> sd\_of\_totalprofit=5\*45

> sd\_of\_totalprofit

[1] 225

> # A)Specify a Rupee range (centered on the mean) such that it contains 95% probability for the annual profit of the company.

> z\_score=round(qnorm(0.975),2) # for 95%

> z\_score

[1] 1.96

> # CI= mu +- (z-score\*sigma)

> lower\_range=mean\_of\_totalprofit-(z\_score\*sd\_of\_totalprofit)

> lower\_range

[1] 99

> upper\_range=mean\_of\_totalprofit+(z\_score\*sd\_of\_totalprofit)

> upper\_range

[1] 981

> #B)specify the 5th percentile of profit (in Rupees) for the company

> zscore=round(qnorm(0.05),3)

> zscore

[1] -1.645

> percentile=round(mean\_of\_totalprofit+(zscore\*sd\_of\_totalprofit))

> percentile

[1] 170

> 540-(1.645\*225)

[1] 169.875

> #C)Which of the two divisions has a larger probability of making a loss in a given year?

> # loss is when profit < 0

> prob\_profit1=pnorm(0,5,3)

> prob\_profit1

[1] 0.04779035

> prob\_profit2=pnorm(0,7,4)

> prob\_profit2

[1] 0.04005916

**(done in python)**